一种基于性格的微博情感分析模型 PLSTM *

袁婷婷，杨文忠，仲丽君，张志豪，向进勇

(新疆大学 信息科学与工程学院，乌鲁木齐 830046)

摘 要：不同性格的用户所具有的语言表达方式不尽相同，现有的情感分析工作很少考虑到用户的性格，针对此问题，提出一种基于性格的微博情感分析模型 PLSTM。该模型首先采用性格识别规则将微博文本分为五个性格集合和一个通用集合，其次针对每种性格文本集合分别训练出一个情感分类器，最后对六个基本情感分类器进行融合，得出最终的情感极性。实验结果显示 PLSTM 方法的 F1 值可以达到 96.95%，表明 PLSTM 比起以前常用的基准情感分析模型在准确率、召回率、F1 值上都有较大提高。

关键词：情感分析；性格；Word2vec ；长短时记忆网络；分类器融合

中图分类号：TP391.1 doi: 10.19734/j.issn.1001-3695.2018.07.0521

Personality-based Microblog sentiment analysis model plstm

Yuan Tingting, Yang Wenzhong, Zhong Lijun, Zhang Zhihao, Xiang Jinyong

(*College of Information Science & Engineering*, *Xinjiang University*, *Urumqi* 830046, *China*)

Abstract: Users of different personalities have different language expressions. Existing sentiment analysis work rarely considers the personality of the user. To solve this problem, this paper proposes a micro-blog sentiment analysis model based personality PLSTM. The model firstly uses the personality recognition rules to divide the microblog text into five personality sets and a universal set, then train a sentiment classifier for each personality set, and finally integrate six basic sentiment classifiers to obtain the ultimate sentiment polarity. The experimental results show that the F1 value of the PLSTM method can reach 96.95%, which indicates that PLSTM has a higher improvement in accuracy, recall rate and F1 value than the commonly used benchmark sentiment analysis model.

Key words: sentiment analysis; personality; word2vec; long and short memory network; classifier Fusion

0 引言

近几年来，随着互联网技术的快速发展和日渐成熟，网络社交平台受到越来越多的人广泛使用。微博、微信等新兴社交平台可以方便快捷地发布文本、图片、视频，方便用户进行信息交流和意见表达，因而深受广大的用户的喜爱。

微博不仅是人与人之间交流和沟通的一种媒介，也是一种在工作和日常生活中表达个人情感的方式。人们可以在微博上了解最新的各种热门事件，发表自己的看法和了解别人的看法。用户在表达观点、传播思想、抒发个人情感的同时，会产生大量带有个人主观情感特征的信息，这些信息中包含着不同倾向性的情感特征，可能反映了用户的爱好和兴趣，也可能对网络舆情的传播产生巨大的影响。因此，通过对微博言论进行情感分析可以衡量微博用户的偏好和政治站位，通过对微博中言论进行情感分析也可了解人们对现实生活中的一些热点事件的看法、预测未来趋势，若微博言论中出现不利于民族团结的、破坏民族团结的言论，可以及时进行控制、引导舆论向着积极正面的方向发展。

情感分析这个术语是 Nasukawa 等人[1]在 2003 年首次使用的。意见挖掘这个术语是 Dave 等人[2]于 2003 年首次使用的。最初的文本挖掘研究[3,4]着重于从文本中提取事实信息。而近来焦点正在转向意见挖掘，也被称为情感分析。这种转变的驱动因素之一是以评论、博客、微博、社交媒体评论等为主的文本越来越多，这类文本中包含大量主观情感。

Pang 等人[5]在电影评论数据集上利用支持向量机（SVM）、朴素贝叶斯（NB）和最大熵等方法研究了评论文本的情感极性判定，但没有考虑到不同性格用户的评论风格也不同这个问题。Yu 等人[6]运用句子相似性和朴素贝叶斯分类器进行主观文本分类。句子相似性方法是基于主观句与主观句之间的相似程度高于主观句与客观句之间的相似程度的假设。魏韡等人[7]分别从文本粒度、文本类型角度介绍了情感分析的研究进展，并介绍了现有的相关资源，但对具体的方法并没有仔细介绍。杨立公等人[8]以文本粒度为视角，从情感词抽取、语料库和情感词典构建、评价对象与意见持有者分析、篇章级情感分析、实际应用五个方面对文本情感分析文献进行了梳理，并作出必要评述。张林等人[9]提出了基于短评论共现的特征筛选方法来进行特征选择，使用短评论的特征来补充长评论的特征。但没有考虑将用户性格也作为特征的一部分 Poria 等人[10]采用融合语音、视频和文本的多模态线索进行情感分析。刘小明等人[11]使用卷积神经网络来进行特征提取，最后训练出基于深度卷积神经网络的互联网短文本情感分类模型。Li 等人[12]提出采用长短期记忆语言模型（LSTM）来进行情感分析。Chen 等人[13]使用了 BiLSTM 和 CNN 两种神经网络方法来提高情感分析的效果。

收稿日期：2018-07-12；**修回日期：**2018-08-23 **基金项目：**国家自然科学基金资助项目（U1603115，61262087）；国家“973”计划资助项目（2014CB340500）；国家自然科学基金重点项目（U1435215）

作者简介：袁婷婷（1993-），女，陕西旬阳人，硕士研究生，主要研究方向为自然语言处理、信息安全（1192815854@qq.com）；杨文忠（1971-），男，副教授，博士，主要研究方向为网络舆情、情报分析、信息安全、无线传感器网络；仲丽君（1992-），女，硕士研究生，主要研究方向为信息安全；张志豪（1995-），男，硕士研究生，主要研究方向为突发事件预警、信息安全；向进勇（1992-），男，硕士研究生，主要研究方向为自然语言处理.

针对微博的情感分析，何炎祥等人[14]提出了一种情感语义增强的深度学习模型。首先使用词向量表示技术来对微博中常用的表情符号构建一个情感空间的特征表示矩阵，在将矩阵和词向量进行乘积运算得到词义到情感空间的映射，并将映射结果输入到一个 MCNN（multi-channel convolution neural network）模型，训练得到一个微博情感分类器。但没有考虑到性格会影响到用户的表达方式，不同性格的用户在表达过程中使用的词语或者表情符号会有所不同。Sun 人[15]等采用基于卷积扩展特征的深度神经网络来进行中文微博的情感分析，实验结果表明，在适当的结构和参数条件下，提出的深度信念网络对情感分类的性能要优于 SVM、NB 等模型。

在性格预测方面，有许多心理学和计算科学研究都探讨了大五模型[16]中人们的语言使用和人格特征之间的关系[17]。大多数研究利用皮尔逊相关系数或斯皮尔曼的秩相关系数来衡量相关性的强度，并识别用户生成内容中不同人格特征相关的重要语言线索[17]。人们也已经开始使用机器学习技术来预测社交媒体中用户的性格特征[20]。Golbeck 等人[20]提取 Twitter 使用特征，结构特征和语言特征，并应用两种回归算法预测用户的人格特征。Bai 等人[21]建议利用多任务回归和增量回归来通过新浪微博（weibo.com）用户的在线行为来预测用户性格。Nowson 等人[22]应用机器翻译模型来解决基于文本的性格预测中的多语言问题。

以上的情感分析和性格预测工作基本上都是不同研究领域的工作。情感分析没有考虑到不同用户的性格的表达情感方式不同，也没有考虑情感分析与性格分析的结合方式。心理学研究表明，性格会影响人们的写作和说话方式，具有相同性格的人倾向于选择类似的情感表达。针对这一问题，本文提出一种基于性格的微博情感分析模型 PLSTM（personality long short term memory network，基于性格的长短时记忆网络）。

1　相关概念

1.1　性格模型

心理学上已经提出了一些性格模型，如 Big Five 模型[16]和 MBTI 模型[23]。在性格模型中，大五模型是较为权威的性格模型[17]，并在心理学和人工智能中被广泛采用[19]。大五模型从五个维度来描述人的性格，即开放性，责任性，外向性，愉悦性和神经质。开放度高的人富有想象力、创造力和好奇心。责任性反映了一个人自律和为机会做好充分准备的程度。高度责任性的人热衷于工作，渴望取得成就。外向性高的人表示喜欢与人交往，而内向者喜欢独处。愉悦性的人是慷慨的、值得信赖的，他们总是愿意去帮助别人。神经质反映了人的情绪的不稳定程度。

1.2　Word2vec 简介

Word2vec 是 Google 推出的一种训练词向量模型的工具，采用的语言模型主要有 CBOW（continuous bag-of-words，即连续的词袋模型）和 skip-gram 两种。Word2vec 可以把文本进行数字化表示即将文本转换为计算机能识别和理解、并进行处理的形式。它是一种神经网络，它可以在使用深度学习算法之前对文本进行预处理。虽然 word2vec 本身没有实现深度学习，但是它把文本变成了深度学习能够理解的向量形式[11]。Word2vec 可以通过训练把文本内容的处理简化为 K 维向量空间中的向量运算，文本语义上的相似度可以通过向量空间上的相似度来进行表示。因此通过 word2vec 训练的词向量可以被很多 NLP 相关的工作使用。例如词性分析、聚类和找同义词等。Word2vec 可以对其词向量进行高效的加法组合运算，Mikolov 在文献[24]中提出的一个优化的单机版本一天可训练上千亿词。

1.3　LSTM 简介

在文本情感分析的研究中，单词之间的顺序关系是至关重要的。Mikolov[25]提出了一种被称为递归神经网络（RNN）的语言模型，该模型被公认为非常适合处理文本序列数据。从理论上讲，RNN 语言模型可以覆盖全文的时间顺序结构，处理长期依赖性问题。然而在实践中，RNN 可能无法成功地学习到相关的知识。当文本的相对信息与要预测的当前位置之间的间隔变大时，由于随时间向后传播算法（BPTT）的后向传播层数过多，会导致训练时的历史信息丢失和梯度衰减或爆炸[26]。为了克服这一困难，Hochreite 等人[27]提出了长短期记忆网络（Long Short Term Memory，LSTM），在某些应用场景下可以得到较好的实验结果。LSTM 通过详细的设计来避免长期依赖，在实践中记住长期信息是 LSTM 的默认行为。目前，LSTM 网络是应用最为广泛的一种，它用 LSTM 单元代替隐藏层中的 RNN 节点，用于保存文本历史信息。LSTM 使用三个门来控制文本历史信息的使用和更新，这些信息分别是输入门，忘记门和输出门。存储单元和三个门设计使 LSTM 能够读取，保存和更新远程历史信息，图 1 显示了 LSTM 的结构。

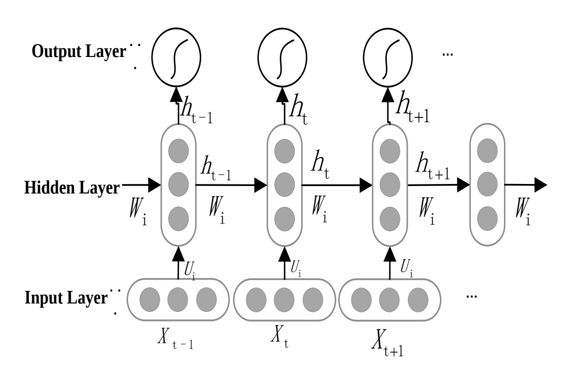


图 1　LSTM 结构

Fig. 1　LSTM structure

2　基于性格的微博情感分析模型

为了进一步提升情感分析的效果，本文提出了一个基于性格的微博情感分析模型。由于性格相同的用户发布的微博文本中往往包含了类似的情感特征，因此首先根据性格将微博文本分到不同的性格集合中。然后针对每组数据集合中的微博文本，采用 word2vec 得到文本中词的向量表示，进而得到微博文本的词向量矩阵，采用该向量矩阵作为 LSTM 的输入，训练出一个情感分类器。最后，为了整合所有基于个性和普通情感分类器的结果，采用集成学习，并构建了一个集成分类器。当对微博文本进行分类时，每个分类器生成输出，然后将其用作集成分类器的输入以产生最终的分类结果。基于性格的情感分类模型框架如图 2 所示，其中 C、A、E 分别指外向性，愉悦性和责任型三个维度，H、L 各性格值得高低，例如 HA 表示高愉悦性、LA 表示低愉悦性。

2.1　性格分类

为了准确地将微博文本分配给不同的集合，需要准确地预测文本具有的性格特色。目前性格预测中主要考虑大五模型中的三个性格维度：外向性（extraversion），愉悦性（agreeableness）和责任性（conscientiousness）。另外两个维度，开放性和神经质，在以前的研究基础上进行预测难度较

大[19,20]因此本文中没有考虑。根据性格的得分值一个性格又分为高、低两个维度，因此本文主要研究高责任性、高愉悦性、低愉悦性、高外向性、低外向性五个维度。由于微博中低责任性的文本太少，本文中暂时没有考虑低责任性这一性格。

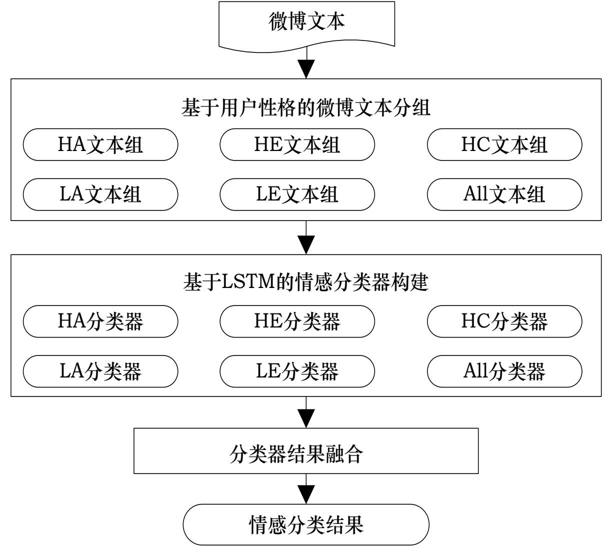


图 2　基于性格的情感分类模型框架

Fig. 2　Personality-based sentiment classification model framework

本文采用的是基于规则的性格分类方法来预测外向性（extraversion），愉悦性（agreeableness）和责任性（conscientiousness）这三个维度的性格。

首先针对每个性格分别建立了一个性格词典，包含了该性格下的常用词语。词典被用来判断一个文本是否具有某类性格的表达。表 1 给出了各性格词典中部分词语。

表 1　性格词典示例

Table 1　Example of personality dictionary

性格	词典包含词语示例
HC	成就、梦想、坚持、奋斗、平等
HA	相信、很棒、爱你、原谅、同意、可怜
LA	愚蠢、勾结、伤害、活该、傻子
HE	哈哈、开心、喜欢、一起、快速、热闹
LE	难过、伤心、孤独、慢慢来、安静

每个性格群体的文本特征反映了相应情感表达的共性。由表 1 可以看出责任性的表达常常是对成就的看法（如努力和失败）。而愉悦性的表达往往与爱情和赞美有关（例如“爱你”、“真棒”），表达更多的是同情（如“悲伤”）。相比之下，低愉悦性程度的表达通常包括指责或辱骂其他人（如“都怪你”和“愚蠢”）。外向性的表达喜欢直接表达积极（例如“开心”）或消极（如“难过”）的情绪。

例如对于文本 X，如果文本中高责任性（HC）表达的词的数量（HC_Cword）或者表情符号的数量（HC_Cemoction）较高，则推断该文本责任性值比较高。表 2 给出了性格预测的主要文本特征。虽然目前篇章级情感分析取得不错的成果，但对于大多数文文档来说一篇文档不止包含一种情感，单从篇章级来进行情感分析分类的效果不如句子级的情感分析。

表 3 是性格判定的规则（其中 p1，p2，…，p10 是规则使用的阈值，阈值的大小由实验决定）。文本满足哪一个性格判定的规则就被划分到该性格集合中，一个文本可以符合多个性格判定规则，因此一个文本也可以同时属于多个性格集合。

表 2 性格预测的文本特征

Table 2　Text characteristics of personality prediction

特征符号	特征符号含义
HC_Cword	文本中高责任性词语的数量
HC_Cemoction	文本中表达高责任性的表情符号的数量
HA_Cword	文本中高愉悦性词语的数量
HA_Cemoction	文本中表达高愉悦性的表情符号的数量
LA_Cword	文本中低愉悦性词语的数量
LA_Cemoction	文本中表达低愉悦性的表情符号的数量
HE_Cword	文本中高外向性词语的数量
HE_Cemoction	文本中表达高外向性的表情符号的数量
LE_Cword	文本中低外向性词语的数量
LE_Cemoction	文本中表达低外向性的表情符号的数量

2.2　情感分类器的结果融合

采用已标记的数据集来构造每个性格集合的基本情感分类器。由于微博文本是词为单位的序列数据，词和词具有长程依赖性，尤其是反映情感和性格方面的词，长程依赖的可能性更大。而 LSTM 是一个序列模型，对于输入数据也是一个序列，并且 LSTM 可以解决长程依赖问题。因此在这里采用的是 LSTM 方法来构造情感分类器。

通用的情感分类器，在训练的过程中与其他常用的特征混合在一起时，那些较少使用的与个性相关的特征可能不起作用。因此，本文选择融合不同基本分类器的分类结果。一个用户可能包含多种性格，因此一个微博文本也可能属于多个性格集合。

首先通过 LSTM 针对每个性格数据集分别训练得到一个情感分类器，在进行预测时分别使用每一个情感分类器对微博文本进行情感倾向的预测，之后对所有六个基本分类器输出的结果进行结果融合。情感分类器的融合过程如图 3 所示。

给定一组微博文本 $(t_1,t_2,\ldots,t_n)$，使用六个 LSTM 基本情感分类器为每条微博文本产生输出 p_{ij}^{+}、p_{ij}^{-}，其中 p_{ij}^{+} 和 p_{ij}^{-} 分别表示第 j 个分类器计算得到的微博文本的积极、消极的概率。基于每个基本情感分类器的输出，通过融合方法将结果进行融合得到最终的情感极性。本文考虑了求和、加权和、中位数三种融合方法。三种融合方法公式如下：

$$l' = \arg\max_i \left(\frac{1}{6}(p_{ij} + p_{ij} + \ldots + p_{ij})\right), i = 1,2,\ldots,c;\ j = 1,2,\ldots 6 \tag{1}$$

$$l' = \arg\max_i (q_1 p_{ij} + q_2 p_{ij} + \ldots + q_6 p_{ij}), i = 1,2,\ldots,c;\ j = 1,2,\ldots,6 \tag{2}$$

$$\mathrm{l}' = \arg\max_{i}(\underset{j}{median}\ p_{ij}), i = 1,2,\ldots,c;\ j = 1,2,\ldots 6 \tag{3}$$

其中: l' 代表最终的情感极性；i 是情感类别数目(i =1,2,3...C)，本文中 i 为 2，情感类别有积极和消极两类； p_{ij} 分别为六个分类器的输出概率分数； $q_1,q_2,\ldots,q_6$ 是各基本分类器的权重。

3　微博情感分析实验

为了验证本文提出的方法的可行性设计了以下实验。本实验的实验环境：操作系统为 Win7，处理器为 Intel Core i5，内存 8 GB，CPU 为 2.5 GHz，开发工具为 PyCharm Community Edition 3.3。

3.1　实验数据

本文的训练数据来源于文献[28]，包括 2009 年 10 月 21 日至 2014 年 12 月 15 日期间新浪微博用户发布的文本。文本包含其微博内容以及作者基本信息，但不包括转发的文本。该数据集包含了 10 474 条文本，并对每条文本都进行了情感极性的标注。其中积极的有 7 562 条，消极的 2 912 条。

测试数据是来自 2012 年 CCF 自然语言处理与中文计算会议提供的中文微博情感分析评测数据，共取 1 100 条文本，其中积极文本 500 条，消极文本 600 条。

表 3　性格判定规则

Table 3　Personality determination rules

规则名称	规则	规则含义
高责任性判定规则	IF *HC*_ Cword ≥ *p1* ∨ HC_Cemoction ≥ *p2* THEN *C* = HC	当文本中包含高责任性词典中的词语数量超过 p1，或者文本中包含高责任性表情符号的数量超过 p2，则判定该文本属于高责任性。
高愉悦性判定规则	IF *LA*_ Cword ≥ *p5* ∨ LA_Cemoction ≥ *p6* THEN *A* = LA	当文本中包含高愉悦性词典中的词语数量超过 p3，或者文本中包含高愉悦性表情符号的数量超过 p4，则判定该文本属于高愉悦性。
低愉悦性判定规则	IF *LA*_ Cword ≥ *p5* ∨ LA_Cemoction ≥ *p6* THEN *A* = LA	当文本中包含低愉悦性词典中的词语数量超过 p5，或者文本中包含低愉悦性表情符号的数量超过 p6，则判定该文本属于低愉悦性。
高外向性判定规则	IF *HE*_ Cword ≥ *p7* ∨ HE_Cemoction ≥ *p8* THEN *E*= HE	当文本中包含高外向性词典中的词语数量超过 p7，或者文本中包含高外向性表情符号的数量超过 p8，则判定该文本属于高外向性。
低外向性判定规则	IF *LE*_ Cword ≥ *p9* ∨ LE_Cemoction ≥ *p10* THEN *E* = LE	当文本中包含低外向性词典中的词语数量超过 p9，或者文本中包含低外向性表情符号的数量超过 p10，则判定该文本属于低外向性。

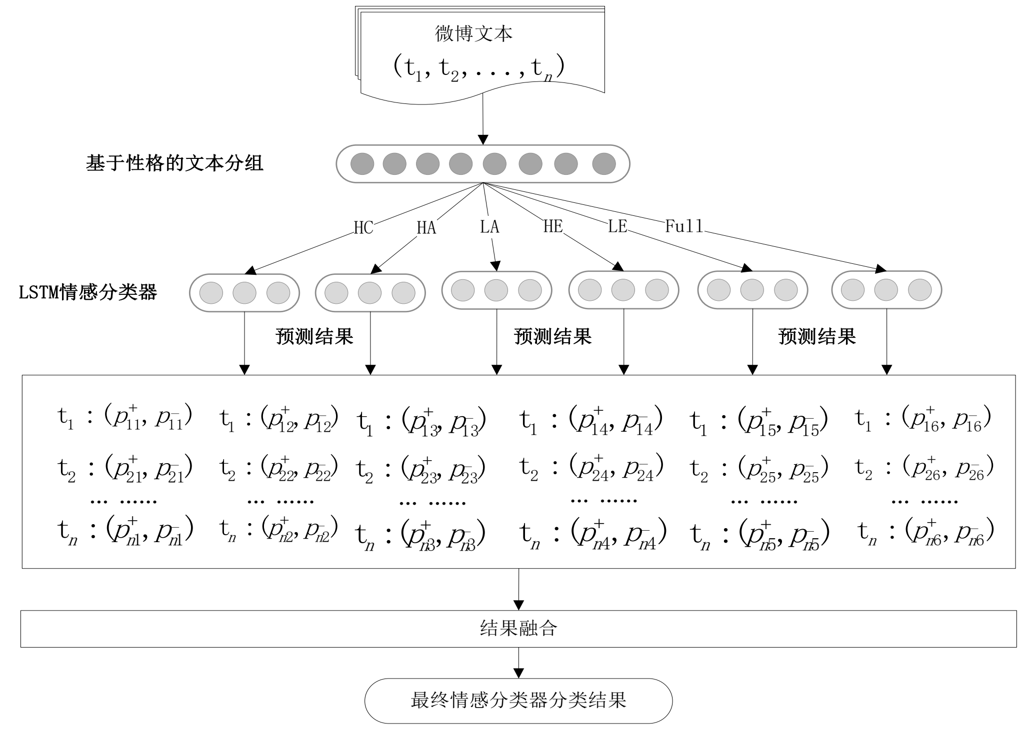


图 3　情感分类器的整合过程

Fig. 3　The process of integration of sentiment classifiers

3.2　基本情感分类器比较

本实验主要用来验证各基本情感分类器的分类准确性，包括五个性格基本情感分类器和一个通用基本情感分类器。训练数据中包括通用数据集合 10474 条，高责任性数据集合（HC）3151 条，高愉悦性数据集合（HA）3188 条，低愉悦性数据集合（LA）3204 条，高外向性数据集（HE）5585，低外向性数据集合（LE）3154 条。实验结果如表 4 所示。

表 4 基本情感分类器比较

Table 4　Comparison of basic sentiment classifiers

性格	ALL	HA	HC	HE	LA	LE
F1 值	88.31	**89.19**	**89.59**	87.39	72.28	83.25

由表 4 可以看出五个性格基本情感分类器中 HA 分类器和 HC 分类器的 F1 值要高于 ALL 分类器，这表明了针对不同性格集合来进行情感分类是有效的。

测试数据中 HA 性格的文本数据和 HC 性格的文本数据最多，LA 性格的文本数据最少。从五个性格基本情感分类器的比较结果可以看出 HA 分类器和 HC 分类器都有较高的 F1 值，而 LA 分类器则具有最低的 F1 值，这也说明某一个性格的文本数据较多时，针对该性格的基本情感分类器是更为有效，说明针对不同性格集合来进行有针对性的情感分类是有意义的。

为了进一步研究每个性格对情感分类最终结果的影响，在进行结果融合的时候去掉一个用户性格集合的情感分类结

果，然后将其性能和 PLSTM 方法得到的结果进行比较。表 5 给出了结果对比，其中 PLSTM-HC、PLSTM-HA、PLSTM-LA、PLSTM-HE、PLSTM-LE 分别代表进行结果融合的时候没有采用 HC、HA、LA、HE 或者 LE 性格文本集合时训练的基于性格的情感分类模型。

从表 5 的结果可以看出在这几种情况下 PLSTM 方法的 F1 值是最高的，说明少了任何一个性格文本集合情感极性分类效果都有下降，表明每种性格文本集合都可以为最终的分类准确性做出贡献，即基于性格的情感分类器的训练应该充分利用各种性格的文本集合，这样训练的集成情感分类器其分类效果较好。

表 5 缺少某一性格文本集合的 plstm 模型实验结果

Table 5: PLSTM model experimental results lacking a set of personality texts

	PLSTM	PLSTM-HA	PLSTM-HC	PLSTM-HE	PLSTM-LA	PLSTM-LE
准确率	96.91	94.45	95.18	95.18	96.91	96.18
召回率	97.00	95.00	94.40	94.80	97.00	96.00
F-Score	**96.95**	94.72	94.79	94.99	96.95	96.09

3.3 融合方法比较

本实验用来验证和选择不同基本分类器的融合方法，主要考虑求和、加权和、中位数三种融合方法。

求和方法中每一个分类器的权重是相等的，即 $q_i = 1/6$；加权和方法中通过交叉验证的方法确定实验中各性格集合的权重。

加权和方法的权重确定采用交叉验证，主要制定标准是准确率尽可能高。这里采用两组数据进行实验。数据 1 是来自 2012 年 CCF 自然语言处理与中文计算会议提供的中文微博情感分析评测数据，共取 1100 条文本，其中积极文本 500 条，消极文本 600 条；数据 2 是在微博上爬取的文本，共有 1500 条，积极和消极的各有 750 条；

由于要得到最优的各个基本情感分类的权重计算工作量太大，本文做了三十组实验取使得准确率最好的一组权重作为近似最优结果。最终得到的各性格集合的权重分别为 $q_1 = 0.16, q_2 = 0.23, q_3 = 0.16, q_4 = 0.16, q_5 = 014, q_6 = 0.15$。表 6 列出了部分权重设置下的准确率对比。

表 6 部分权重设置对比

Table 6 Partial weight setting comparison

权重（ q_1,q_2,q_3,q_4,q_5,q_6 ）	数据 1	数据 2
0.2,0.16,0.16,0.16,0.16,0.16	96.42	85.67
0.25,0.15,0.15,0.15,0.15,0.15	96.27	84.87
0.16,0.23,0.16,0.16,014,0.15	**96.91**	**85.86**
0.15,0.27,0.15,0.14,0.15,0.14	96.55	83.86
0.15,0.15,0.23,0.16,0.15,0.16	96.81	85.66
0.14,0.16,0.25,0.15,0.15,0.15	96.81	83.25
0.16,0.16,0.16,0.21,0.15,0.16	96.43	85.40
0.15,0.15,0.15,0.24,0.15,0.16	96.63	84.26
0.16,0.16,0.15,0.16,0.21,0.16	96.41	84.32
0.14,0.16,0.15,0.16,0.24,0.15	96.36	85.2
0.14,0.15,0.15,0.15,0.16,0.25	96.45	85.06
0.14,0.15,0.14,0.15,0.14,0.28	96.27	85.46

数据 2 是从微博爬取下来的数据，在预处理之后进行人工标注，存在一定的误差和主观性，因此所得出的实验结果比数据 1 的结果低很多。

表 6 中的结果显示在两组数据中都是在 HC 的权重最高的情况下取得了近似的最优结果。而权重通常与分类器在训练集上的准确率成正比，即准确率越高，权重越大[29]。从表 4 结果中可以看出五个基本分类器中 HC 的准确率是最高的，所以其权重 q_2 也相应的应该大一些。这与从表 6 所得出的结果一致。

三种融合方法的实验结果如表 7 所示。

表 7 融合方法比较

Table 7 Comparison of fusion methods

方融合法	准确率	召回率 R	F1 值
求和	96.58	97.00	96.78
加权和	96.91	97.00	**96.95**
中位数	95.35	96.23	95.79

由表 7 可以看出三种方法中加权和方法的 F1 值最高。其中求和方法是假设所有的分类器都有同等权重，但在实际应用中，不同的分类器其重要程度不同所具有的权重也应该有所不同，因此其效果要比加权和方法差。中位数方法是将几个分类器结果按大小排序，选出居于中间位置的数，在本文中这个方法的效果也比加权和方法差。因此本文选择加权和方法作为最终的融合方法。

3.4 对比实验

为了验证本文提出的 PLSTM 模型的有效性，将 PLSTM 模型和其他基准情感分类模型 SVM、LSTM、Kim[30]在 2014 提出的 CNN-rand、CNN-static、CNN-non-static 在准确率、召回率、F1 值等方面做了对比实验室，实验结果见表 8。

表 8 实验结果

Table 8 Experimental results

方法	准确率 P	召回率 R	F1 值
SVM	58.91	87.6	70.45
CNN-rand	73.03	85.37	78.72
CNN-static	73.52	88.49	80.31
CNN-non-static	73.55	88.36	81.79
LSTM	88.82	87.80	88.31
PLSTM	**96.91**	**97.00**	**96.95**

从表 8 的实验结果可以看出，本文中提出方法的准确率比其他几种基准方法都要高，证明了本文提出方法的有效性。由于具有相同性格的用户其表达方式趋于一致，针对每个性格文本集合进行训练得出的情感分类器比起通用的情感分类器更具有针对性，同时在最后采用了集成学习来对每个性格的情感分类器和一个通用情感分类器进行结果融合，这样可以使得那些较少使用的与个性相关的特征也起到作用。因此本文提出的方法取得了比其他基准情感分类器都要好的效果。

4 结束语

本文针对不同性格的用户所具有的语言表达方式不尽相同，现有的情感分析工作很少考虑到用户的性格这一问题提出一种基于性格的微博情感分析 PLSTM 模型。文章主要分析了现有情感和性格分析的现状和存在的问题，给出了 PLSTM 模型的整体架构，通过实验验证了 PLSTM 模型的有效性。

PLSTM 只使用了基于 LSTM 的分类器来进行情感分类器的训练，在后续工作中将考虑尝试使用双向长短期记忆网络 BiLSTM、深度信念网络 DNN 等深度神经网络模型，看是否能取得更好的结果，同时可以考虑把另外两种性格开放性和神经质也加入进来进一步提高基于性格的情感分析模型

PLSM 的效果。

参考文献：

[1] Nasukawa T, Yi J. Sentiment analysis: capturing favorability using natural language processing [C]//Proc of International Conference on Knowledge Capture. 2003:70-77.

[2] Dave K, Pennock D M, Pennock D M. Mining the peanut gallery: opinion extraction and semantic classification of product reviews [C]//Proc of International Conference on World Wide Web. ACM, 2003:519-528.

[3] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation [J]. Journal of Machine Learning Research, 2012(3):993-1022.

[4] Hofmann T. Probabilistic latent semantic indexing [C]// Proc of International ACM SIGIR Conference on Research and Development in Information Retrieval.New York:ACM Press, 1999:50-57.

[5] Bo P, Lee L, Vaithyanathan S. Thumbs up: sentiment classification using machine learning techniques [C]// Proc of Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics.2002:79-86.

[6] Yu Hong. Towards answering opinion questions: separating facts from opinions and identifying the polarity of opinion sentences [J]. Pediatrics, 2003, 116(3):58-59.

[7] 魏韡,向阳, 陈千.中文文本情感分析综述[J].计算机应用,2011,31(12):3321-3323. (Wei Wei, Xiang Yang, Chen Qian.Survey on Chinese text sentiment analysis [J]. Journal of Computer Applications 2011,31(12):3321-3323.)

[8] 杨立公,朱俭,汤世平.文本情感分析综述[J]. 计算机应用 2013,33(6):1574-1607. (Yang Ligong, Zhu Jian, Tang Shiping. Survey of text sentiment analysis [J]. Journal of Computer Applications 2013,33(6):1574-1607.)

[9] 张林，钱冠群，樊卫国,等. 轻型评论的情感分析研究[J]. 软件学报, 2014,25(12):2790-2807. (Zhang Lin, Qian Guanqun, Fan Weiguo, et al. Emotion analysis of light comment [J] .Journal of Software, 2014,25(12): 2790-2807.)

[10] Poria S, Cambria E, Howard N, et al. Fusing audio, visual and textual clues for sentiment analysis from multimodal content[J]. Neurocomputing, 2016, 174:50-59.

[11] 刘小明,张英,郑秋生. 基于卷积神经网络模型的互联网短文本情感分类[J]. 计算机与现代化,2017(4):73-77. (Liu Xiaoming, Hang Ying, Zheng Qiusheng. Sentiment classification of short texts on internet based on convolutional neural networks model [J]. CAM 2017(4):73-77.)

[12] Li Dan, Jiang Qian. Text sentiment analysis based on long short-term memory [C]//Proc of IEEE International Conference on Computer Communication and the Internet. 2016:471-475.

[13] Chen Tao, Xu Ruifeng, He Yulan, et al. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN [J]. Expert Systems with Applications, 2017, 72:221-230

[14] 何炎祥，孙松涛，牛菲菲,等. 用于微博情感分析的一种情感语义增强的深度学习模型[J]. 计算机学报，2017,40(4):773-790. (He Yanxiang, Sun Songtao, Niu Feifei, et al. A deep sentiment learning model for sentiment analysis of Weibo [J]. Chinese Journal of Computers, 2017, 40 (4): 773-790.)

[15] Sun Xiao, Li Chengcheng, Ren Fuji. Sentiment analysis for Chinese microblog based on deep neural networks with convolutional extension features [J]. Neurocomputing, 2016, 210:227-236.

[16] Digman J M. Personality Structure: Emergence of the five-factor model [J]. Psychology, 1990, 41(1):417-440.

[17] Hirsh J B, Peterson J B. Personality and language use in self-narratives [J]. Journal of Research in Personality, 2009, 43(3):524-527.

[18] Yarkoni T. Personality in 100,000 words: a large-scale analysis of personality and word use among bloggers [J]. Journal of Research in Personality, 2010, 44(3):363-373.

[19] Qiu L, Lin H, Ramsay J, et al. You are what you tweet: Personality expression and perception on Twitter [J]. Journal of Research in Personality,2012,46(6):710-718.

[20] Golbeck J, Robles C, Edmondson M, et al. Predicting Personality from Twitter [C]//Proc of IEEE Third International Conference on Privacy, Security, Risk and Trust. 2012:149-156.

[21] Bai S, Hao B, Li A, et al. Predicting Big Five Personality Traits of Microblog Users [C]// Proc of IEEE/WIC/ACM International Joint Conferences on Web Intelligence. 2013:501-508.

[22] Nowson S, Perez J, Brun C, et al. XRCE personal language analytics engine for multilingual author profiling notebook for PAN at CLEF 2015 [EB/OL]. 2015. http://ceur-ws.org/Vol-1391/100-CR.pdf.

[23] Myers I B, Mccaulley M H, Most R. Manual: a guide to the development and use of the Myers-Briggs type indicator [M]// Manual: a guide to the development and use of the Myers-Briggs type indicator. Consulting Psychologists Press, 1985.

[24] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality [C]//Advances in Neural Information Processing Systems. 2013:3111-3119.

[25] Mikolov T. Recurrent neural network based language model[C]//Proc of Conference of the International Speech Communication Association.2010:1045-1048.

[26] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult [J]. IEEE Trans on Neural Networks, 1994, 5(2):157-166.

[27] Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural Computation, 1997, 9(8):1735.

[28] Lin Junjie, Mao Wenji, Zeng D D. Personality-based refinement for sentiment classification in microblog [J]. Knowledge-Based Systems, 2017, 132:204-214

[29] 寿质彬. 基于神经网络模型融合的图像识别研究[D]. 广州:华南理工大学,2015. (Shou Zhibin. Research on Image Recognition Based on Neural Network Model Fusion [D]. Guangzhou: South China University of Technology,2015.)

[30] Kim Y. Convolutional Neural Networks for Sentence Classification [J]. Eprint Arxiv, 2014

[31] Etworks for Sentence Classification [J]. Eprint Arxiv, 2014